

Classificação de patologias laríngeas por meio de características do espaço de fase reconstruído

Pinho, P. H. U., Couras, M. F. K. B., Dantas, E. C. S., Costa, S. C. e Correia, S.E.N.

Resumo— Este artigo trata da classificação entre vozes normais e vozes afetadas por patologias laríngeas (edema de Reinke e paralisia nas pregas vocais), utilizando informações das imagens do espaço de fase reconstruído dos sinais de voz. São empregadas sete características obtidas pelos métodos da contagem de caixas e da diferença, considerando atrasos diferentes e quatro dimensões de caixas: 40x40, 20x20, 10x10 e 5x5. A classificação é realizada com as medidas extraídas individualmente e de forma combinada, através de redes neurais artificiais. A melhor taxa de acurácia, 97,53%, foi obtida na classificação entre vozes normais e vozes afetadas por edema de Reinke.

Palavras-Chave—Reconstrução do Espaço de Fase. Redes Neurais, Patologias Laríngeas, Processamento de Sinais de Voz.

Abstract— This article deals with the classification between normal voices and voices affected by laryngeal pathologies (Reinke's edema and paralysis in vocal cords), using the information from the reconstructed phase space images of the voice signals. Seven characteristics, obtained from the methods of the boxes counting and the difference methods, were employed, considering different delays and four dimensional grids: 40x40, 20x20, 10x10 and 5x5. The classification is made with the measures taken individually and combined, through artificial neural networks. The best accuracy rate, 97.53%, was obtained in the classification between normal voices and voices affected by Reinke's edema.

Keywords— Reconstruction of phase space. Neural networks, Laryngeal Pathologies, voice signal processing.

I. INTRODUÇÃO

A análise da qualidade vocal por meio da análise acústica, empregando técnicas de processamento digital de sinais, vem sendo muito estudada no decorrer dos últimos anos. Essa forma de análise permite conhecer ou prever as condições anatômicas e fisiológicas das estruturas laríngeas [1]. As alterações na laringe têm aumentado consideravelmente devido a práticas como o tabagismo e o alcoolismo, além do abuso vocal. O diagnóstico precoce de patologias laríngeas pode influenciar no tratamento e na cura da doença [2-4].

Entre as medidas acústicas tradicionais podem ser citadas a frequência fundamental e suas medidas de perturbação como *jitter* e *shimmer*, entre outras [3,4]. Pesquisas indicam que a análise não linear pode quantificar padrões que não foram bem avaliados com as medidas de perturbação, concluindo que este tipo de análise pode quantificar padrões aperiódicos apresentados em patologias severas [5, 6].

Os processos periódicos, multi-periódicos ou quase-periódicos podem ser identificados a partir de métodos tradicionais de análises de sinais experimentais como a função de autocorrelação e espectro de potência. No entanto, essas técnicas não permitem a análise de sinais que possuam

dinâmica caótica [7]. A reconstrução do espaço de fase é um dos métodos da análise dinâmica não linear que vêm sendo muito estudado nos últimos anos, pois o espaço de fase mostra padrões visuais de dinâmica vocal que representam consideravelmente componentes caóticos e determinísticos presentes na voz. Obtém-se, dessa forma, uma ferramenta de informação visual que pode ser usada para complementar os métodos utilizados na análise e diagnóstico de patologias na voz [8].

Diversas medidas têm sido propostas para analisar um sinal caótico, baseadas na reconstrução do espaço de fase, combinado com métodos estatísticos, para determinar algumas medidas quantitativas que possam caracterizar o sinal, tais como dimensão de correlação, expoentes de Lyapunov, função de informação mútua e medidas de entropia, que investigam a estabilidade local de um sistema não linear [2,5,9].

Pesquisas utilizam a análise dinâmica não linear, voltadas para a análise de sinais biomédicos, a exemplo de ecocardiogramas, por serem sinais que apresentam um comportamento não linear, possuindo pequenos fenômenos de dessincronização, que é difícil de ser identificado por simples observações, sendo considerado um sinal caótico [10,11].

Amann et al. [12] obtiveram a imagem dos espaços de fase reconstruídos de sinais de eletrocardiograma normais e de taquicardia ventricular e propuseram um método que particiona o espaço de fase em pequenas caixas, classificando os sinais pela contagem do número de caixas visitadas para cada trajetória do sinal.

Roopaei et al. [10] estendem o trabalho de Amann [12], propondo o uso de três novas medidas para avaliar a autossimilaridade de sinais de arritmias cardíacas, as quais baseiam-se na diferença de duas imagens de espaços de fase reconstruídos de um mesmo sinal. Os resultados experimentais obtidos, utilizando uma grade de dimensão 40x40, foram considerados superiores, quando comparados às medidas dimensão de correlação, maior expoente de Lyapunov e entropia aproximada.

Em sinais de voz, diversas técnicas têm sido empregadas para detecção de patologias na laringe, por meio de análise dinâmica não linear, empregando medidas tradicionais do espaço de fase [2,5]. O classificador utilizando redes neurais MLP (Multi-Layer Perceptron) tem sido bastante empregado na classificação de vozes patológicas [4], [13], [14]. No trabalho proposto por Umopathy et al. [13] foram utilizadas características tempo-frequência, em um classificador MLP, obtendo-se uma classificação de 93,4%. No trabalho proposto por Srinivasan et al. [14] é feita uma análise paramétrica com os coeficientes mel cepstrais, classificando os sinais utilizando redes neurais MLP e, comparando o desempenho com as Redes Neurais Probabilísticas e de Regressão, tendo se destacado, com o melhor desempenho, a rede MLP.

Pinho, P. H. U., Couras, M. F. K. B., Dantas, E. C. S., Costa, S. C. e Correia, S.E.N., Departamento de Engenharia Elétrica, Instituto Federal da Paraíba, João Pessoa-PB, Brasil, E-mails: pablohenriqueifpb@gmail.com, kallynna.mary@gmail.com, evandson@live.com, silvana@ifpb.edu.br e suzete@ifpb.edu.br. Este trabalho foi parcialmente financiado pela COPEX e PRPIPG - IFPB.

Neste trabalho, são empregadas medidas da imagem do espaço de fase propostas por [10] e [12], com o intuito de discriminar entre sinais de vozes normais e sinais de vozes afetados por patologias na laringe (edema de Reinke e paralisia nas pregas vocais). Diversos valores de atraso foram considerados para a construção dos espaços de fase dos sinais.

Os métodos da contagem de caixas e da diferença são avaliados, considerando-se quatro grades de dimensões 40x40, 20x20, 10x10 e 5x5. Foi realizada a classificação dos sinais empregando as características de forma individual e combinada, usando uma rede neural MLP [16].

O artigo está organizado da seguinte forma: Na Seção II é descrito o método da reconstrução do espaço de fase, na Seção III é apresentada a descrição da base de dados e da metodologia empregada. Na Seção IV são apresentados os resultados obtidos e discussão, seguida da Seção V, em que são apresentadas as conclusões.

II. RECONSTRUÇÃO DO ESPAÇO DE FASE

O Teorema de Takens [15] demonstra que utilizando o método das coordenadas defasadas é possível reconstruir certas propriedades do espaço de fase a partir de uma série temporal em que vetores m -dimensionais são reconstruídos. [2,12]:

$$\xi_i = \{x(t_i), x(t_i + \tau), \dots, x(t_i + (m - 1)\tau)\} \quad (1)$$

em que $x(t)$ representa a série temporal, m representa a dimensão de imersão e o τ representa o tempo de atraso.

O comportamento do traçado obtido na reconstrução do espaço de fase, usando o τ ótimo, representa a dinâmica vocal. Quanto mais regular o espaço de fase reconstruindo, mais periodicidade o sinal possui [8].

O tempo de atraso ótimo é determinado pelo ponto em que ocorre o primeiro valor mínimo na curva de informação mútua, que representa a autocorrelação da série medindo a dependência do sinal consigo mesmo [16]. Esta curva oferece a melhor estimativa do tempo de atraso ou passo de reconstrução ótimo para o sinal. Este tempo realça a dinâmica aleatória e não determinística do sinal, representado por τ ótimo. Na Figura 1, é apresentado a curva de informação mútua (a) e o espaço de fase reconstruído (b) para um sinal de voz normal. A partir do espaço de fase pode ser observado o comportamento aperiódico ou caótico do sinal [9].

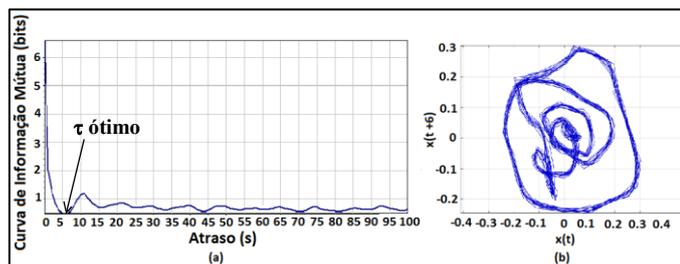


Fig. 1 (a) Curva de Informação Mútua para obtenção do τ ótimo para um sinal normal e (b) Espaço de fase reconstruído para o sinal da curva (a), com τ ótimo =6 e dimensão de imersão $m =2$.

III. MATERIAIS E MÉTODOS

Nesta seção é feita a descrição da base de dados, da metodologia empregada no trabalho.

A. Base de Dados

Foram utilizados 149 sinais extraídos da base de dados *Disordered Voice Database, Modelo 4337*, gravada pelo

Massachusetts Eye and Ear Infirmary (MEEI) [17]. Esta base de dados é amplamente estudada, como pode ser visto nos trabalhos [4, 13, 14]. A base de dados contém pronúncias sustentadas da vogal /a/. Dos sinais selecionados, 53 são de vozes normais, e 96 sinais de vozes patológicas, sendo 44 de sinais de vozes afetadas por edema e 52 afetados por paralisia nas pregas vocais. Os sinais foram amostrados na frequência de 25 kamostras/segundo para sinais patológicos e 50 kamostras/segundo para sinais normais, em que os últimos foram sub-amostrados para 25 kamostras/segundo com o intuito de preservar o tamanho do segmento em amostras para obtenção do espaço de fase. Os sinais utilizados possuem duração média de 1s para os sinais de voz patológicos e 3s para os sinais de voz normais. Estes sinais são previamente processados, não contendo intervalos de silêncio antes ou depois da locução.

B. Metodologia

Inicialmente, é realizada a aquisição do sinal a partir da base de dados. Posteriormente, é extraído o tempo de atraso para cada sinal, para realizar a reconstrução do espaço de fase de cada sinal de voz. Os espaços de fase reconstruídos foram armazenados em imagens com tamanho 680 x 397 pixels. As características foram extraídas através de dois métodos: contagem de caixas e o método da diferença. Esses valores foram armazenados como padrões de referência e foram utilizados pelo classificador para realizar a classificação dos sinais em saudável ou patológico, ou ainda, discriminando a patologia (edema ou paralisia nas pregas vocais).

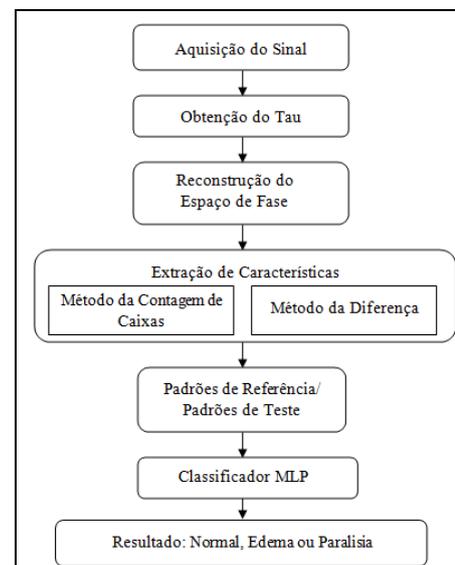


Fig. 2. Diagrama em blocos do sistema de classificação de vozes patológicas.

C. Tempo de Atraso (τ)

O tempo de atraso para a reconstrução do espaço de fase, representando a dinâmica do sistema (τ ótimo) foi encontrado através da curva de informação mútua [16]. Como o valor do τ ótimo varia para cada sinal analisado, foram calculados os valores médio (τ med), mínimo (τ min) e máximo (τ max) para cada classe (normal, edema e paralisia), com os quais foram construídos espaços de fase para cada sinal. Os valores dos tempos de atraso obtidos para cada classe estão apresentados na Tabela I.

TABELA I. VALORES PARA O TEMPO DE ATRASO.

Classe	τ ótimo	τ min	τ med	τ max
Normal	5-10	5	7	10
Edema	1-20	1	9	20
Paralisia	2-34	2	12	34

D. Reconstrução do Espaço de Fase

A reconstrução do espaço de fase foi realizada para 10 ciclos do sinal, em que cada ciclo corresponde a um período de *pitch*. Foi realizada a reconstrução do espaço de fase utilizando os quatro valores obtidos para o tempo de atraso (τ ótimo, τ min, τ med e τ max). Dessa forma, tem-se, para cada sinal, quatro espaços de fase, a partir dos quais serão obtidas as características pelos métodos empregados.

Na Figura 3 são apresentados diagramas de fase de sinais de vozes normais e de vozes afetadas por edema e por paralisia nas pregas vocais, considerando os quatro atrasos de tempo correspondentemente.

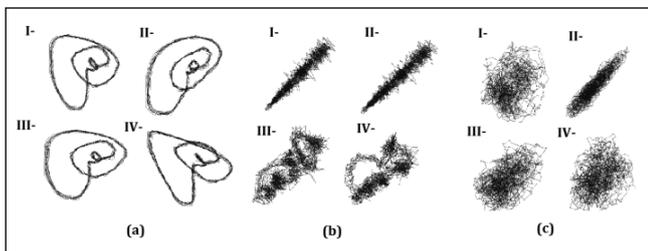


Fig. 3. Espaços de fase reconstruídos para um sinal de voz: (a) normal; (b) edema; (c) paralisia nas pregas vocais (I - τ ótimo; II - τ min; III - τ med e IV - τ max).

E. Método da Contagem de Caixas

No método da contagem de caixas, a imagem do espaço de fase é dividida em caixas, que correspondem a um grupo de pixels. Caso parte da trajetória do espaço de fase esteja contida dentro da caixa, esta é contabilizada como uma caixa ocupada, sendo que ao final do processo é contabilizada a quantidade de caixas visitadas, relacionada ao número total de caixas, de acordo com a Equação (2) [10], [12].

$$d(x) = \frac{n^\circ \text{ de caixas visitadas}}{n^\circ \text{ total de caixas}} \quad (2)$$

em que d estima o grau de complexidade ou a dimensão caótica da trajetória do espaço de fase e x o atraso empregado para a reconstrução do espaço de fase.

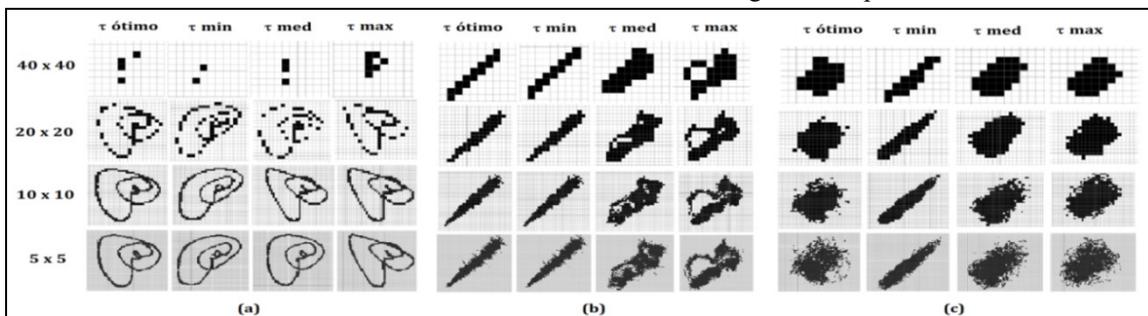


Fig. 4. Espaço de fase reconstruído para sinais de voz: (a) normal; (b) edema (c) paralisia nas pregas vocais com as diversas grades.

Neste trabalho, foram utilizadas quatro grades com caixas de tamanhos diferentes: 40x40, 20x20, 10x10 e 5x5, para observar qual delas obteria melhor desempenho na classificação.

Na Figura 4 estão ilustradas as imagens dos espaços de fase representados na Figura 4 para sinais de voz normal, com edema e com paralisia nas pregas vocais, para as diversas grades empregadas.

F. Método da Diferença

O método procura encontrar a autossimilaridade de um sinal utilizando as duas imagens da reconstrução do espaço de fase com valores de atraso diferentes. É calculada a diferença entre duas trajetórias do mesmo sinal, utilizando as imagens da reconstrução do espaço de fase em tempos diferentes. Quanto mais caótico for o sinal, menor será sua autossimilaridade [10].

A partir do valor de d , obtido anteriormente em cada imagem, pelo método da contagem de caixas, é calculada a diferença entre os valores obtidos a partir das imagens dos espaços de fase, considerando os atrasos τ min, τ med e τ max, apresentados na Tabela I.

Na Figura 5 está ilustrado o método da diferença com grade 10x10 para um sinal de voz normal, um sinal de voz afetado por edema e um sinal de voz afetado por paralisia utilizando a reconstrução do espaço de fase para os valores de τ med e τ min.

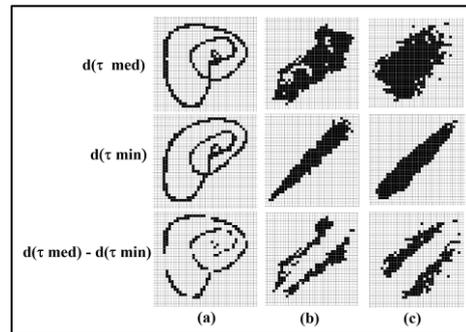


Fig. 5. Método da diferença com a grade 10 aplicada em um sinal de voz (a) normal; (b) afetado por edema; e (c) afetado por paralisia nas pregas vocais.

G. Classificação dos sinais de voz

Foram considerados quatro casos de classificação distintas: (i) Patologia vs. Normal, em que na classe Patologia estão incluídos os sinais afetados por edema e por paralisia; (ii) Edema vs. Normal; (iii) Paralisia vs. Normal; e (iv) Edema vs. Paralisia.

Extraídos os padrões de referências, foi empregada uma rede neural *feedforward*. Na etapa de treinamento, foi utilizado o algoritmo supervisionado do Gradiente Conjugado

Escalonado (SCG) [19], por necessitar de poucas épocas de treinamento e estar presente, por meio de bibliotecas de *software*, em diversas plataformas. Os sinais foram separados,

aleatoriamente, em dois conjuntos distintos, dos quais 2/3 foram utilizados para treinamento e 1/3 para teste.

Para maior confiabilidade, a rede neural foi executada 10 vezes, a fim de realizar a validação cruzada [4], para cada vez foi usada uma divisão aleatória de exemplos para treino/teste diferente, sendo os resultados descritos neste trabalho como os valores médios e seus respectivos desvios padrões.

Para os testes foram realizadas diversas combinações das características de entrada, a fim de encontrar a combinação que pudesse maximizar os resultados. As características de entrada são: (i) $d(\tau \text{ min})$; (ii) $d(\tau \text{ med})$; (iii) $d(\tau \text{ max})$; (iv) $d(\tau \text{ ótimo})$; (v) Diferença absoluta entre $\tau \text{ med}$ e $\tau \text{ min}$ ($d(\tau \text{ med}) - d(\tau \text{ min})$); (vi) Diferença absoluta entre $\tau \text{ max}$ e $\tau \text{ med}$ ($d(\tau \text{ max}) - d(\tau \text{ med})$); (vii) Diferença absoluta entre $\tau \text{ max}$ e $\tau \text{ min}$ ($d(\tau \text{ max}) - d(\tau \text{ min})$).

Na topologia da rede empregada, a quantidade de neurônios na camada de entrada é dependente da quantidade de características selecionadas. Nove neurônios foram empregados na camada oculta e dois neurônios na camada de saída, uma para cada caso de classificação. O neurônio na saída que contiver o maior valor de ativação será usado para identificar a classe.

Foram observadas as medidas de acurácia, sensibilidade e especificidade, em que a acurácia mede a capacidade do classificador de identificar corretamente a presença ou a ausência de patologia. Já a especificidade representa a proporção de pessoas sem a doença, cujo teste dá negativo, indicando o quão bom é o método empregado na identificação dos indivíduos com vozes não patológicas. A sensibilidade representa a proporção de pessoas com a patologia de interesse que têm o resultado do teste positivo, indicando o quão bom é teste para identificar os indivíduos com a patologia.

IV. RESULTADOS

Na classificação dos sinais, foram consideradas diferentes grades. Na Tabela II estão apresentados os resultados, relacionando a medida, de forma individual, e a grade que proporcionaram o melhor desempenho.

Obteve-se a melhor acurácia média na distinção entre vozes afetadas pela patologia edema e vozes normais, utilizando o $d(\tau \text{ min})$, pelo método da contagem de caixas, e para o método da diferença com o $d(\tau \text{ min})$ e o $d(\tau \text{ med})$, ambos com a grade 10x10. Entre as três classes, a pior classificação foi para a distinção entre vozes normais e vozes afetadas por paralisia nas pregas vocais, proporcionando uma acurácia média de 86% para as melhores medidas.

Na Tabela III, é apresentada a classificação com as medidas combinadas. As taxas de acurácia média encontradas são superiores a 94%, com a taxa de especificidade superior a 95% e a sensibilidade até 99%.

A característica $d(\tau \text{ min})$, obtida no método da contagem de caixas, aparece em todas as combinações, associada a outras medidas para a discriminação entre edema vs. normal, paralisia vs. normal e patologia vs. normal.

Na distinção entre patologias e vozes normais observa-se, pela Tabela III, que a melhor combinação utilizou quatro características para distinção entre vozes afetadas por edema e vozes normais, com acurácia média de 97% e sensibilidade atingindo 99%. Sendo que, em todos os resultados, a melhor grade foi a 10x10. Apenas para discriminação entre vozes afetadas por paralisia e vozes normais a melhor combinação foi utilizando a grade 5x5. Na discriminação entre as patologias, obteve-se acurácia próxima a 80%, com destaque para a grade 10x10.

TABELA II. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS INDIVIDUAIS.

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	88,39 ± 0,890	90,38 ± 1,272	87,29 ± 1,728	$d(\tau \text{ med}) - d(\tau \text{ min})$	10x10
Edema vs. Normal	90,82 ± 0,285	93,21 ± 0,308	87,95 ± 0,485	$d(\tau \text{ min})$	10x10
	90,72 ± 1,640	94,91 ± 0,846	85,68 ± 4,136	$d(\tau \text{ med}) - d(\tau \text{ min})$	10x10
Paralisia vs. Normal	86,48 ± 0,397	90,38 ± 1,175	82,50 ± 0,449	$d(\tau \text{ med}) - d(\tau \text{ min})$	5x5
	86,10 ± 0,806	88,87 ± 1,937	83,27 ± 0,705	$d(\tau \text{ med}) - d(\tau \text{ min})$	10x10
Edema vs. Paralisia	65,90 ± 0,421	62,45 ± 2,969	69,42 ± 2,944	$d(\tau \text{ max}) - d(\tau \text{ min})$	20 x 20

TABELA III. RESULTADO PARA A CLASSIFICAÇÃO COM AS CARACTERÍSTICAS COMBINADAS.

Classificação	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	Medida	Grade
Patologia vs. Normal	94,70 ± 0,463	93,21 ± 0,898	95,52 ± 0,313	$d(\tau \text{ ótimo})$, $d(\tau \text{ min})$, $d(\tau \text{ med})$	10x10
Edema vs. Normal	97,53 ± 0,168	99,25 ± 0,308	95,45	$d(\tau \text{ min})$, $d(\tau \text{ med}) - d(\tau \text{ min})$, $d(\tau \text{ max}) - d(\tau \text{ med})$, $d(\tau \text{ max}) - d(\tau \text{ min})$	10x10
Paralisia vs. Normal	95,71 ± 0,670	97,74 ± 0,548	93,65 ± 0,995	$d(\tau \text{ ótimo})$, $d(\tau \text{ min})$, $d(\tau \text{ max})$, $d(\tau \text{ max}) - d(\tau \text{ min})$	5x5
Edema vs. Paralisia	79,62 ± 1,146	82,45 ± 1,567	76,73 ± 1,911	$d(\tau \text{ ótimo})$, $d(\tau \text{ min})$, $d(\tau \text{ max})$, $d(\tau \text{ med}) - d(\tau \text{ max})$	10x10

Na Figura 6 estão ilustrados os boxplots dos melhores resultados para o método da contagem de caixas com $d(\tau \text{ min})$ e para o método da diferença com $d(\tau \text{ med})$ e $d(\tau \text{ min})$ utilizando a grade 10×10 , ilustrando os valores da média e do desvio encontrado para cada classe.

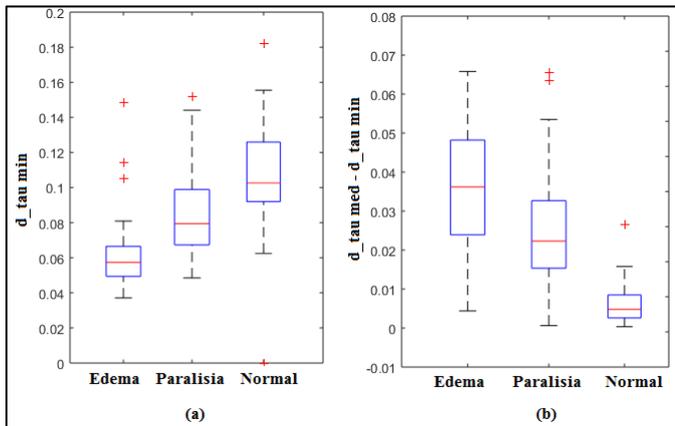


Fig. 6. Boxplots das melhores medidas para: (a) método da contagem de caixas; e (b) método da diferença.

Observa-se que, para os sinais normais, o $d(\tau \text{ min})$ apresenta os maiores valores, tendo em vista que o espaço de fase reconstruído (Figura 3) ocupa uma área mais ampla do que os patológicos que tendem a se alinhar na diagonal, dado que o valor de $\tau \text{ min}$ para esses sinais são menores (Tabela I). Com relação à diferença entre $d(\tau \text{ med})$ e $d(\tau \text{ min})$, há uma maior dispersão entre os valores.

V. CONCLUSÕES

Os resultados obtidos mostram que o uso de características extraídas do espaço de fase para diagnosticar a presença de patologias nas pregas vocais é bastante interessante e promissor.

As medidas realizadas com o τ ótimo para a contagem de caixas não se mostraram eficientes na separação dos diferentes tipos de classes abordadas. Observa-se também que o tamanho da grade influencia no resultado, sendo que a grade 10×10 foi a que obteve melhor desempenho, comparada às outras grades.

Quanto aos métodos, observa-se que o método da diferença foi mais eficiente para discriminar as classes do que o método da contagem de caixas, apresentando acurácia média acima de 94%, na discriminação entre vozes normais e patológicas. Na discriminação entre patologias (Edema vs. Paralisia), o desempenho foi inferior ao que foi obtido entre vozes saudáveis e patológicas. Com o intuito de melhorar o desempenho na classificação, pretende-se investigar o uso de outros classificadores, a exemplo de Máquinas de Vetor de Suporte (*Support Vector Machine* – SVM), além de adicionar à base de dados, outros tipos de patologias na laringe.

AGRADECIMENTOS

Os autores agradecem a PRPIPG e a COPEX do Instituto Federal de Educação Ciência e Tecnologia da Paraíba, Campus João Pessoa, pelo financiamento parcial da pesquisa e ao PPGEE pelo apoio científico.

REFERÊNCIAS

[1] I. C. Zwetsch, “Processamento do sinal de voz para o diagnóstico diferencial de alterações laringeas”, dissertação de mestrado, Faculdade de

Engenharia, Pontifícia Universidade Católica de Porto Alegre, RS, 2006. Disponível em: <http://hdl.handle.net/10923/3192>.

[2] W. C. de A. Costa, S. C. Costa, F. M. de Assis e B. G. Aguiar Neto, “Classificação de sinais de vozes saudáveis e patológicas por meio da combinação entre medidas da análise dinâmica não linear e codificação combinatória linear”, *Rev. Brasileira de Engenharia Biomédica*, vol. 29, nº 1, pp. 3-14, Mar, 2013 [Online]. Disponível em: <http://dx.doi.org/10.4322/rbeb.2013.010>.

[3] W. C. de A. Costa, F. M. Assis, B. G. Aguiar Neto, S. C. Costa e V. J. D. Vieira, “Pathological Voice Classification Based on Recurrence Quantification Measures”, apresentado no XXX Simpósio Brasileiro de Telecomunicações- SBrT 12. Set. 2012. [Online]. Disponível em: http://sbrt.org.br/sbrt2012/publicacoes/98742_1.pdf.

[4] V. J. D. Vieira, “Avaliação de Distúrbios da Voz por meio de Análise de Quantificação de Recorrência”, dissertação de mestrado, Instituto Federal da Paraíba, PB, 2014. Disponível em: <http://www.ifpb.edu.br/campi/joao-pessoa/cursos/pos-graduacao-stricto-sensu/mestrado-em-engenharia-eletrica/imagem/Avaliacao%20de%20Disturbios%20da%20Voz%20por%20meio%20de%20Analise%20de%20Quantificacao%20de%20Recorrencia%20%20Dissertacao%20de%20Vinicius.pdf>.

[5] M. E. Dajer, “Análise de sinais de voz com padrões visuais de dinâmica vocal”, tese de doutorado, Escola de Engenharia de São Carlos, Universidade de São Paulo, SP, 2010. [Online]. Disponível em: www.teses.usp.br/teses/disponiveis/82/82131/tde-16052006-095323/pt-br.php.

[6] J. Lee, Seong-Hee Choi, J. J. Jiang, Min-Soo Hahn e Hong-Shik Choi, “Perturbation and Nonlinear Dynamic Analysis of Sustained Vowels in Normal and Pathological Voices”, *Journal of the Korean society of speech science*, vol. 2, nº 1, pp. 113-120, Fev. 2010.

[7] W. C. de A. Costa, “Análise Dinâmica Não Linear de Sinais de Voz para Detecção de Patologias Laringeas”, tese de doutorado, Universidade Federal de Campina Grande, Paraíba, PB, 2012.

[8] D. G. Galdino, “Análise Acústica não Linear dos Padrões Visuais de Dinâmica Vocal (PVDV) de Homens Adultos”, dissertação de mestrado, Escola de Engenharia de São Carlos, Instituto de Química de São Carlos da Universidade de São Paulo, SP, 2012.

[9] R. Behroozmand; F. Almasjanj; M. H. Moradi, “Pathological Assessment of Vocal Fold Nodules and Polyp Using Accoustic Perturbation and Phase Space Features”, apresentado na IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, vol. 2, Mai. 2006. [Online]. Disponível em: 10.1109/ICASSP.2006.1660528.

[10] M. Roopaei, R. Boostani, R. R. Sarvestani, M. A. Taghavi e Z. Azimifar, “Chaotic based reconstructed phase space features for detecting ventricular fibrillation”, *Biomedical Signal Processing and Control*, vol. 5, pp. 318-327, Out. 2010. [Online]. Disponível em: doi:10.1016/j.bspc.2010.05.003.

[11] G.; Das, S. Koulaouzidis, G. Capiello, E. B. Mazomenos, K. Maharatna e J. Morgan, “A novel approach for the diagnosis of ventricular tachycardia based on phase space reconstruction of ECG”, *International Journal of Cardiology*, vol. 172, nº 1, pp. 31-33, Mar. 2014. [Online]. Disponível em: <http://dx.doi.org/10.1016/j.ijcard.2013.12.088>.

[12] A. Amann, R. Tratnig e K. Unterkofler, “Detecting Ventricular Fibrillation by Time-Delay Methods”, apresentado na IEEE Transactions on Biomedical Engineering, vol. 54, pp. 174-177, Jan. 2007. [Online]. Disponível em: 10.1109/TBME.2006.880909.

[13] K. Umopathy, S. Krishnan, V. Parsa e D. G. Jamieson, “Discrimination of Pathological Voices Using a Time-Frequency Approach”, apresentado na IEEE Transactions on Biomedical Engineering, vol. 52, nº 3, Mar 2005. [Online]. Disponível em: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1396382.

[14] V. Srinivasan, V. Ramalingam e P. Arulmozhi, “Artificial Neural Network based Pathological Voice Classification using MFCC Features”, *International Journal of Science*, Vol. 3, nº 1, pp. 291 – 302, 2014. [Online]. Disponível em: <http://www.ijset.net/journal/255.pdf>.

[15] F. Tankens, “Detecting strange attractores in turbulence,” *Dynamical systems and turbulence*, Warwick 1980, pp. 366-381, 1981.

[16] A. M. Fraser e H. L. Swinney, “Independent coordinates for strange attractors from mutual information”, *Phys Rev.Let.*, vol. 33, 1134-1140, 1986.

[17] Kay Elemetrics. *Disordered Voce Database. Model 4337*. 3th ed. United Satates, 1994.

[18] S. Haykin. “Redes Neurais, Princípios e Práticas”. Porto Alegre, RS. Bookman, 2001.

[19] A. P. Silva, T. B. Ludermer e L. M. Almeida, (2011), “Método AGCRN-CR para busca automática de redes neurais artificiais”, ENIA – Encontro Nacional de Inteligência Artificial, Sessão Poste, Artigo 7. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/enia/2011/0052.pdf>.